2021-01-25\_lacre

Gonzalo García-Castro

10/11/2021

## Background

Infants learning two languages face the challenge of having to learn two distinct sets of words. Some of the words across both languages refer to similar concepts: for instance, *chair* and *silla* (in Spanish). These pairs of words are called *translation equivalents*. Sometimes, translation equivalents also overlap in form, frequently due to their shared etymological origin. These form-similar translation equivalents, such as *boot* and *bota* (in Spanish), are referred to as *cognates*.

There is extensive literature on how cognateness impacts the dynamics of lexical access during words comprehension and production. One of its main findings is that bilinguals process cognates (such as *boot* and *bota*) more differently than non-cognates (such as *chair* and *silla*). For instance, Costa, Caramazza, and Sebastian-Galles (2000) asked Spanish-Catalan bilinguals to name a series of familiar pictures in Spanish. Part of the pictures referred to objects whose labels in Spanish and Catalan were cognates, whereas the other pictures referred to non-cognates. On average, participants named cognate pictures faster than non-cognate pictures. This effect was independent of lexical frequency and, critically, was not observed when Spanish monolinguals performed the same task. This suggests that bilinguals activated labels in both languages, and that their phonology interacted to facilitate naming.

Later studies provided evidence that the locus of this facilitation effect was at the lexical level, and not just at the articulatory level [reference needed here]: this effect has also been reported during comprehension, at both the visual-orthographic and the auditory-phonology modalities) (e.g., Thierry and Wu 2007). Overall, these findings have led researchers to hypothesise that lexical access is language-non selective, that is, bilinguals activate lexical representations in both languages in parallel during comprehension and production, based on their semantic and phonological overlap.

This cross-talk between languages has also been suggested to be present in children’s and toddlers’ lexicon (Von Holzen and Mani 2012; Bosma and Nota 2020), but it is still unclear how it impacts the trajectories of language acquisition at these early ages. In this study we sought to investigate the developmental trajectories of lexical access in a cohort of Catalan-Spanish and English-Spanish bilinguals.

To do this, we used an adaptation of Mani and Plunkett (2010)’s priming paradigm. This paradigm consists of a spoken word recognition task in which participants are first presented with a picture in silence in the centre of the screen for 1.5 seconds. Then, two other pictures are presented side-by-side during 2 seconds, and one of them is named. The other is not. Looking times to the target picture are recorded from name onset. From now on, I will refer to the silent picture presented first as the *prime*, to the named picture as the *target*, and to the non-named picture presented next to the target as *distractor*.

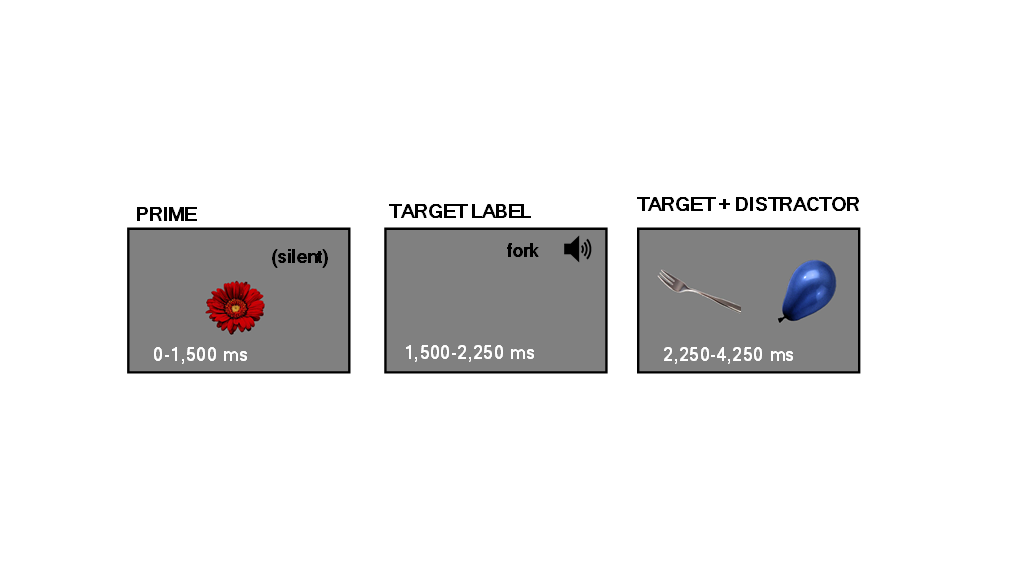
In their experiment, Mani and Plunkett manipulated the phonological overlap between the prime label and the target label: in some trials, both labels shared phonological onset (*bee*-*ball*), whereas in other trials they did not share phonological onset (*cow*-*book*). In all trials, the distractor label did not share onset with either of the prime or target labels (e.g., *teeth*). The authors registered 18 month-old participants’ looking preference for the target object after name onset, as a proxy of word recognition, and found that target preference was stronger after a phonologically related prime than after a phonologically unrelated prime. There results have three main implications:

1. Phonologically related prime pictures facilitated target recognition
2. This is only possible if participants lexicalised the prime pictures, even though this picture was presented in silence
3. The label participants generated upon prime picture presentation was phonologically detailed, and this phonological form interacted with the auditory recognition of a subsequently presented word

In a subsequent study, the authors found the inverse effect in 24 month-old toddlers: target looking preference was weaker after a phonologically related prime picture compared to after an unrelated one (Mani and Plunkett 2011). The authors suggested that the change in the direction of the effect observed between 18 and 24 month-old participants was due to the increment in vocabulary size and the establishment of (inhibitory) links between phonologically related lexical entries that might have occurred between both ages. Supporting this account, the inhibition effect they reported was stronger as the number of phonological neighbours of the target increased: the recognition of words surrounded by more phonologically related neighbours suffered the most from the priming effect.

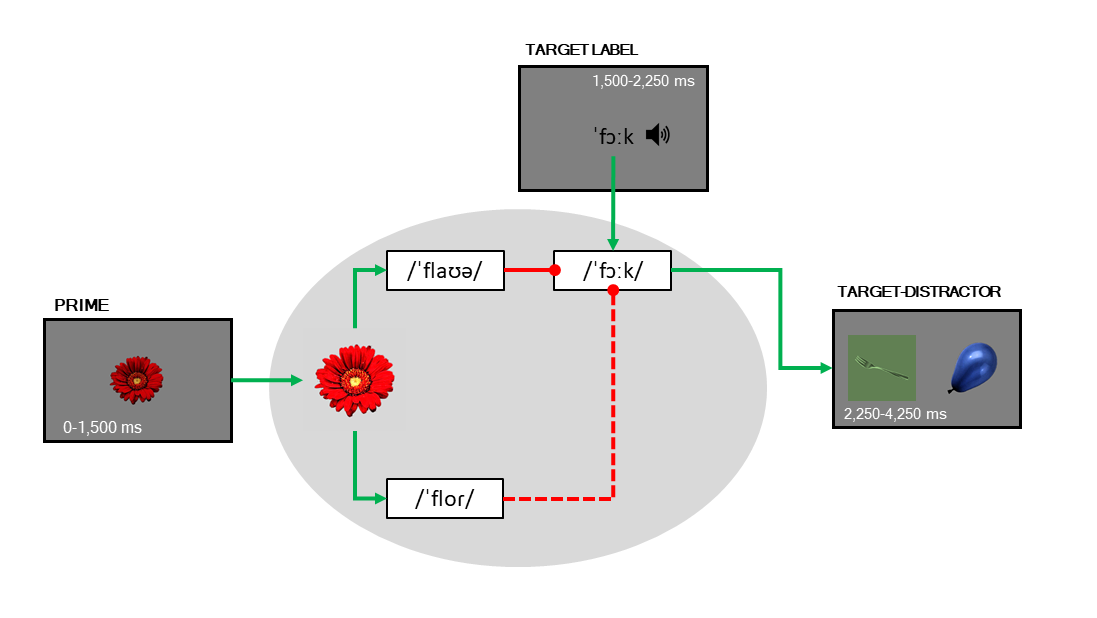
We adapted this paradigm to test the language-non selectivity account in bilingual toddlers. Following Mani & Plunkett’s original implementation, In our task half of the prime-target pairs in the trial lists shared phonological onset, and the other half did not share phonological onset. We added a novel manipulation: half of the primes in the phonologically related condition were cognates and the other half were non-cognates. We classified two translation equivalents as cognates if they shared phonological onset. Here is an illustration of each trial type:

This is what a trial looks like [see figure below]:



## Hypotheses

If participants lexicalise the prime pictures in a language-non selective way, they should generate two labels upon prime picture presentation: one in each language. Therefore, if lexical access is language non-selective, in cognate trials two prime labels should interfere with target recognition (since both prime labels share onset with the target). In non-cognate trials, only one prime label should interfere with the target. Finally, in unrelated trials no prime labels should interfere with the target. As a result, we expect bilingual participants’ target looking preference to be weaker in cognate trials than in non-cognate trials, and in cognate and non-cognate trials than in unrelated trials.



We also collected data from same-aged monolingual participants as a control group: in this group we also expect to find weaker target looking in cognate and non-cognate trials, compared to unrelated trials, but no differences between cognate and non-cognate trials (these participants should be insensitive to the cognate status of the primes).

## Participants

We planned to collect data from participants learning English and/or Spanish in Oxford, and from participants learning Catalan and/or Spanish in Barcelona. Due to limitations imposed by the pandemic and lockdown, data collection in Oxford was severely interrupted, so I will present preliminary data from participants tested in Barcelona exclusively.

We tested participants at three age points: 21, 25, and 30 months, to explore their trajectories of lexical access through these sensitive ages in which participants’ vocabulary size is growing a full speed [Reference here.]

We classified participants as monolinguals or bilinguals by estimating their overall exposure to each of their languages using Bosch and Sebastián-Gallés (2001). Following the consensus reached by the ManyBabies consortium (Byers-Heinlein et al. 2021), participants exposed to any languages more than 80% of the time were classified as monolinguals. Otherwise, participants were classified as bilinguals. Participants exposed more than 10% to a third language or a language other than Catalan and Spanish (in Barcelona) or English and Spanish (in Oxford) were excluded from data analysis.

This is how testing sessions are distributed across ages and language profiles:

These 156 testing sessions correspond to 106 distinct participants. Of these participants, 1, 57 only participated once (I will refer top these participants as cross-sectional participants), 2, 41 participated twice, and 3, 8 participated three times.

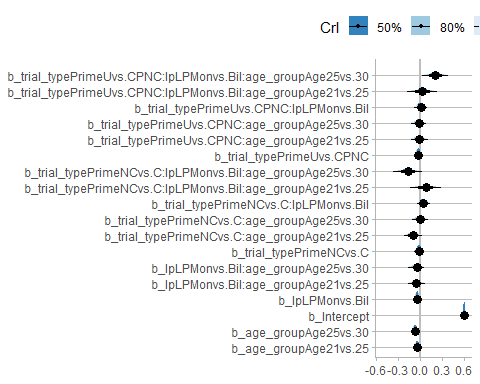
## Data analysis

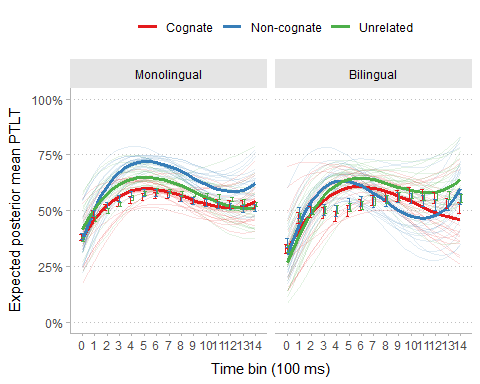
In each testing session, participants were presented with 32 trials (8 cognate, 8 non-cognate and 16 unrelated). In each trial, we used an eye-tracker to register participants’ looking time to the target picture. We delimited a time window of interest between 250 ms after the target onset and 150 ms before the target offset. This resulted in a 1,500 ms long time window. We divided this time window into 15 time bins of 100 ms each, and calculated the adjusted logit of the samples in which participants’ gaze was located in the target coordinates, out of the total of valid samples in that time bin. This resulted in 15 observations per participant, per age group, and per trial. Our final dataset comprised 77166 data points from 4273 distinct trials: 1051, cognate trials, 1043 trials, and 2115 trials.

To model these data we used Growth Curve Analysis (Mirman 2017). First, we included time\_bin as a fixed effect in our model to account for the correlation between the data points collected within each trial. We included this variable as a third degree polynomial, in order to account for the possibly non-linearity of the target fixations across the same trial (target fixations might increase rapidly at the beginning of the trial and decrease at the end of it. We then included our predictors of interest: age group (age\_group), language profile (lp), trial type (trial\_type), and the two-way interaction between language profile and trial type. Finally, we added random intercepts and slopes by participant and by target picture, to account for the possible correlation between time series from the same participant or trial (more details of the model at the end).

We estimated our model using the Bayesian framework, under which we estimate the probability of each value of the sampling space of each parameter, given the data we have observed. to do this, we used the brms R package (Bürkner 2017).

## Results





## References

Bosch, Laura, and Núria Sebastián-Gallés. 2001. “Evidence of Early Language Discrimination Abilities in Infants from Bilingual Environments.” *Infancy* 2 (1): 29–49.

Bosma, Evelyn, and Naomi Nota. 2020. “Cognate Facilitation in Frisian–Dutch Bilingual Children’s Sentence Reading: An Eye-Tracking Study.” *Journal of Experimental Child Psychology* 189: 104699.

Bürkner, Paul-Christian. 2017. “Brms: An r Package for Bayesian Multilevel Models Using Stan.” *Journal of Statistical Software* 80 (1): 1–28.

Byers-Heinlein, Krista, Angeline Sin Mei Tsui, Christina Bergmann, Alexis K Black, Anna Brown, Maria Julia Carbajal, Samantha Durrant, et al. 2021. “A Multilab Study of Bilingual Infants: Exploring the Preference for Infant-Directed Speech.” *Advances in Methods and Practices in Psychological Science* 4 (1): 2515245920974622.

Costa, Albert, Alfonso Caramazza, and Nuria Sebastian-Galles. 2000. “The Cognate Facilitation Effect: Implications for Models of Lexical Access.” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26 (5): 1283.

Mani, Nivedita, and Kim Plunkett. 2010. “In the Infant’s Mind’s Ear: Evidence for Implicit Naming in 18-Month-Olds.” *Psychological Science* 21 (7): 908–13.

———. 2011. “Phonological Priming and Cohort Effects in Toddlers.” *Cognition* 121 (2): 196–206.

Mirman, Daniel. 2017. *Growth Curve Analysis and Visualization Using r*. Chapman; Hall/CRC.

Thierry, Guillaume, and Yan Jing Wu. 2007. “Brain Potentials Reveal Unconscious Translation During Foreign-Language Comprehension.” *Proceedings of the National Academy of Sciences* 104 (30): 12530–35.

Von Holzen, Katie, and Nivedita Mani. 2012. “Language Nonselective Lexical Access in Bilingual Toddlers.” *Journal of Experimental Child Psychology* 113 (4): 569–86.